CMPSC 465 Fall 2023

Data Structures & Algorithms
Antonio Blanca and Yana Safonova

Quiz 1 Solutions

Monday, Aug 28, 2023

1. (2 pts.) The implementation of Insertion Sort from lecture (for sorting in ascending order) always runs in linear time for input lists with which property?

(a) Sorted in ascending order.

(b) No duplicate elements.

(c) Sorted in descending order.

(d) A list with the first $n-100$ elements sorted in ascending order.

(e) A list with the first $n-\lceil\sqrt{n}\rceil$ elements sorted in ascending order.

Answer (a), (d), or (a) and (d) are all acceptable for points. However, the true solution is (a) and (d). A sorted list requires no swapping so it runs in linear time. Similarly, in the case of (d), only $O(1)$ of the elements require $O(n)$ swaps; the overall running time in this case would be linear. Note that for (e), the running time could be $O(n^{3/2})$.

2. (2 pts.) While performing insertion sort on the array $\{7, 5, 4, 2, 3, 1\}$, which of the following is NOT a transition state of the array:

(a) 4 5 7 2 3 1

(b) 2 3 4 5 7 1

(c) 7 5 4 2 1 3

(d) 5 7 4 2 3 1

(e) None of the above

Answer (c). Insertion sort runs from index $k = 1$ to $k = n-1$ and ensures that after the k-th iteration the first $k+1$ elements are sorted.

3. (2 pts.) When analyzing an algorithm, the most important metric is the algorithm's best-case performance.

(a) True

(b) False

Answer False. Best-case inputs are often very fast, but also very rare. Worst-case analysis provides a more useful guarantee.

4. (2 pts.) $f(n) = 3n^2$ and $g(n) = 2n^3$. Which of the following is correct?

I $f(n) = O(g(n))$
II $g(n) = O(f(n))$

(a) I is correct

(b) II is correct

(c) Both I and II are correct

Answer (a) As $n^2 < n^3$ in terms of order of growth.

5. (2 pts.) Consider the following algorithm to identify if a given natural number n is prime. Let $T(n)$ be the number of times the for loop is executed by the program on input n. Which of the following is true:

```
for i := 2 to √n do
    if n mod i == 0 then
        print "Not Prime"
        return 0
    end
end
return 1
```

(a) $T(n) = O(\sqrt{n})$ and $T(n) = \Omega(\sqrt{n})$

(b) $T(n) = O(n)$ and $T(n) = \Omega(\sqrt{n})$

(c) $T(n) = O(\sqrt{n})$ and $T(n) = \Omega(1)$

(d) $T(n) = O(n)$ and $T(n) = \Omega(n)$

Answer (c). The for loop runs a maximum of $\sqrt{n}$ times and a minimum of 1 time. Note that this is not the same as a best-case or worst-case analysis. For the lower bound, observe that for some values of n (e.g., even numbers), it only takes on iteration; therefore, $\Omega(\sqrt{n})$ is not a valid lower bound in general.

CMPSC 465 Fall 2023

Data Structures & Algorithms
Antonio Blanca and Yana Safonova

Quiz 2 Solutions

Monday, Sep 11, 2023

1. (2 pts.) Any non-root node can simply be deleted from a binary min heap without any extra steps needed.

- ◯ True
- ◯ False

Answer False
To delete a node, its value is swapped with the value of the last node on the heap, and then we many need to run Heapify-UP or Heapify-Down.

2. (2 pts.) The array $[1,5,3,8,4,6,7]$ is a min heap.

- ◯ True
- ◯ False

Answer False
This is not a min heap, but it is almost a min heap. The incorrect value is 4, which is the right child of 5. In a min heap, no value can be less than its parent.

3. (2 pts.) It is possible to compute all the Fibonacci numbers up to the n-th one with an algorithm whose running time is $O(n^2)$.

- ◯ True
- ◯ False

Answer True
Storing each Fibonacci number in an array and calculating the n-th requires $O(n)$ iterations. Because the Fibonacci numbers grow exponentially, the n-th has $O(n)$ bits, which makes each addition $O(n)$ and the total loop $O(n^2)$.

4. (2 pts.) Heapify-up and Heapify-down each may require $O(\log n)$ swaps, where n is the number of items in the heap.

- ◯ True
- ◯ False

Answer True
Proof derived in class. In short, the maximum number of swaps required is the height of the tree. Because a heap is a full binary tree, the height of the tree is $O(\log n)$.

5. (2 pts.) With an array representation of a binary heap, how do you get the position of the parent of a node at position p?

- ◯ $2p+1$
- ◯ $\lceil \frac{p+1}{2} \rceil$
- ◯ $2(p+1)$
- ◯ $\lfloor \frac{p-1}{2} \rfloor$

Answer $\lfloor \frac{p-1}{2} \rfloor$
Division by 2 is the central operation to find the parent. The -1 and the floor operation ensure the left and right children find the same parent.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 3

Recitation Section: **Date: Monday, Sep 18, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) We can apply the Master Theorem to recurrences of the form $T(n) = a \cdot T(\frac{n}{b}) + O(2^{n \cdot d})$ (where $a > 0$, $b > 1$ and $d \geq 0$) and conclude that:

$$T(n) = \begin{cases} \Theta(n^d) & \text{if } d > \log_b a \\ \Theta(n^d \log n) & \text{if } d = \log_b a \\ \Theta(n^{\log_b a}) & \text{if } d < \log_b a \end{cases}$$

- ◯ True
- ◯ False

Answer False
Master Theorem applies to recurrences of the form $T(n) = a \cdot T(\frac{n}{b}) + O(n^d)$. Everything besides the given $T(n)$ is correct.

2. (2 pts.) What is the tightest correct upper bound to the recurrence relation $T(n) = 4 \cdot T(\frac{n}{2}) + O(n)$?

- ◯ $O(n)$
- ◯ $O(n^2)$
- ◯ $O(n \log n)$
- ◯ $O(2^n)$

Answer $O(n^2)$
Plugging parameters into the Master Theorem, we have, $a = 4$, $b = 2$, $d = 1$. Thus we get $O(n^{\log_2 4}) = O(n^2)$.

3. (2 pts.) Given two arrays of numbers $x = [5, 6, 10]$ and $y = [1, 3, 7]$. What would be the result of $Merge(x, y)$ in the merge-sort algorithm?

- ◯ $[1, 3, 6, 5, 7, 10]$
- ◯ $[1, 7, 5, 3, 6, 10]$
- ◯ $[1, 3, 5, 6, 7, 10]$
- ◯ $[3, 5, 6, 1, 7, 10]$

Answer $[1, 3, 5, 6, 7, 10]$
When Merge runs on two sorted arrays, it always returns a sorted array.

4. (2 pts.) There is an $O(n^{1.59})$ algorithm for multiplying two n-bit integers. (Note: $O(\log_2 3) \approx 1.59$)

- ◯ True
- ◯ False

Answer True
An $O(n^{1.59})$ multiplication algorithm was shown in class.

5. (2 pts.) Suppose you are given the following two algorithms:

- Algorithm A solves problems by dividing them into four subproblems of half the size, recursively solving each subproblem, and then combining the solutions in linear time.
- Algorithm B solves problems of size n by dividing them into nine subproblems of size $n/3$ recursively solving each subproblem, and then combining the solutions in $O(n^2)$ time.

Which of these algorithms has a faster running time?

- ◯ A
- ◯ B

Answer A
Applying the Master Theorem to both algorithms shows that A runs in $O(n^2)$ time, while B runs in $O(n^2 \log n)$

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 4

Recitation Section: **Date: Monday, Sep 25, 2023**

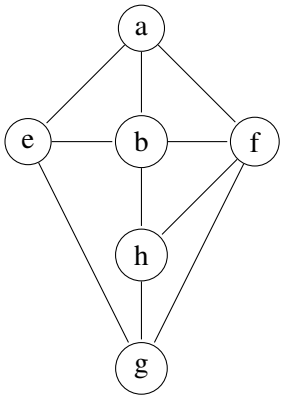
Student Name: **PSU Email ID:**

1. (2 pts.) We have a graph with n vertices, with its edges stored in an $n \times n$ adjacency matrix. How much time will it take to verify that a single edge exists using the array?

- ◯ $O(n \log n)$
- ◯ $O(n)$
- ◯ $O(1)$
- ◯ $O(n^2)$

Answer $O(1)$
An adjacency list is stored as an array. The lookup time for a single element is $O(1)$.

2. (2 pts.) Consider the following graph,



Which of the following are plausible DFS traversals of the graphs? **Mark all correct answers**.

- ☐ a b e g h f
- ☐ a b f e h g
- ☐ a b f h g e
- ☐ a f b e h g

Answer "a b e g h f" and "a b f h g e"
DFS explores all available neighbors before backtracking.

3. (2 pts.) What is the number of edges present in an undirected complete graph having n vertices?

- ◯ $\frac{n(n+1)}{2}$
- ◯ $\frac{n(n-1)}{2}$
- ◯ $n(n-1)$

Answer $\frac{n(n-1)}{2}$
Number of pairs of n vertices is $\binom{n}{2} = \frac{n(n-1)}{2}$.

4. (2 pts.) The structure of a directed graph determines which edges are tree edges and which are cross edges in its DFS tree. In other words, the identity of the tree edges and cross edges is independent of the DFS implementation used.

- ◯ True
- ◯ False

Answer False
DFS can choose to visit the neighbors of the current node in any order. This ordering decides which edges are tree edges and which are cross edges.

5. (2 pts.) In a DFS tree, back edges are defined as:

- ◯ Edges that lead to the root node.
- ◯ Edges that point into the current node.
- ◯ Edges that lead to a child node.
- ◯ Edges that lead to an ancestor node.

Answer Edges that lead to an ancestor node.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 5

Recitation Section: **Date: Monday, Oct 2, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) A back edge (u,v), $(u \to v)$ of the DFS algorithm on a directed graph satisfies which of the following conditions?

- ◯ $\text{pre}[u] < \text{pre}[v] < \text{post}[v] < \text{post}[u]$
- ◯ $\text{pre}[v] < \text{pre}[u] < \text{post}[u] < \text{post}[v]$
- ◯ $\text{pre}[v] < \text{post}[v] < \text{pre}[u] < \text{post}[u]$
- ◯ None of the above

Answer $\text{pre}[v] < \text{pre}[u] < \text{post}[u] < \text{post}[v]$
A back edge implies v is an ancestor of u, so v is visited first, then u, then u finishes, then v finishes.

2. (2 pts.) Suppose in a directed graph, A and B are strongly connected components, and there is an edge from a vertex in A to a vertex in B. Then:

- ◯ $max_{w \in A}\ post(w) > max_{v \in B}\ post(v)$
- ◯ $max_{w \in A}\ pre(w) < max_{v \in B}\ pre(v)$
- ◯ $max_{w \in A}\ post(w) < max_{v \in B}\ post(v)$
- ◯ None of the above

Answer $max_{w \in A}\ post(w) > max_{v \in B}\ post(v)$
When running DFS, if A is visited first, the vertex with an edge to B will have a higher post number than every vertex in B. If B is visited first, A cannot be reachable, or they would not be separate SCCs. DFS will assign post numbers to B without visiting A, and will have to visit A later.

3. (2 pts.) The node that receives the lowest *post number* in a depth-first search must lie in a sink strongly connected component.

- ◯ True
- ◯ False

Answer False.
Consider a metagraph with three SCCs $B \to A \to C$. If DFS begins in A, a vertex in A may be assigned a post number before the edge to C is taken.

4. (2 pts.) Finding a topological ordering requires sorting the vertices by post number, which is an $O(|V|\log|V|)$ operation. Therefore, finding a topological ordering cannot be done faster than $O(|V|\log|V|+E)$ time.

- ◯ True
- ◯ False

Answer False.
The post numbers are assigned in sorted order, so it can be done in $O(|V|+|E|)$.

5. (2 pts.) Consider an undirected graph G. Let T be a DFS traversal forest (union of trees). Let u be a vertex in G and v be the first vertex visited after visiting u. Which of the following statements is always true?

- ◯ $\{u,v\}$ must be an edge in G, and u is a descendant of v in a tree in T
- ◯ $\{u,v\}$ must be an edge in G, and v is a descendant of u in a tree in T
- ◯ If $\{u,v\}$ is not an edge in G, then u and v have the same parent in a tree in T
- ◯ If $\{u,v\}$ is not an edge in G, then u is a leaf in a tree in T

Answer If $\{u,v\}$ is not an edge in G, then u is a leaf in a tree in T.
In other words, u may have no unvisited neighbors, causing DFS to backtrack and visit v.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 6

Recitation Section: **Date: Monday, Oct 9, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) Dijkstra's Algorithm cannot be applied on:

- ◯ Directed weighted graphs
- ◯ Graphs having negative weights
- ◯ Undirected unweighted graphs

Answer Graphs having negative weights
Negative edge weights allow paths shorter than the first one to visit a node.

2. (2 pts.) In a weighted graph where the weights of all edges are unique, there is always a unique shortest path from a source to a destination.

- ◯ True
- ◯ False

Answer False
A counterexample is a graph where one edge of weight 3 connects the source and destination, and another path through a third node has weights 1 and 2.

3. (2 pts.) What is the maximum number of times the Decrease Key operation would be performed during Dijkstra's Algorithm?

- ◯ $|V|$
- ◯ $|E|$
- ◯ $|V|+|E|$
- ◯ $|V|-1$

Answer $|E|$
At most, every edge discovers a new shortest path to its endpoint.

4. (2 pts.) Given a directed graph where the weight of every edge is the same and positive, we can most efficiently find the shortest path from a given source to every destination using?

- ◯ BFS
- ◯ Dijkstra's Algorithm
- ◯ DFS
- ◯ None of the above

Answer BFS
A node's level in the BFS tree is directly related to the distance from the root. For positive weights, just multiply the depth by the shared weight. Note that Dijkstra will return a correct answer, but it is slower than BFS.

The original wording of this problem accidentally allowed negative edges, which cannot be so easily handled with BFS. Due to the confusion this may have caused, we'll give everyone points for question 4.

5. (2 pts.) In Dijkstra's Algorithm, we can use a min-heap to retrieve the node with the next smallest distance value in each iteration. This optimization improves the running time of from $O(|V|^2)$ to $O(|V|\log|V|)$.

- ◯ True
- ◯ False

Answer False
We can use a min-heap to optimize Dijkstra's Algorithm, but the improved running time is only $O((|V|+|E|)\log|V|)$. As discussed in question 3, we might have to run Decrease Key $O(|E|)$ times, which adds $O(|E|\log|V|)$ to the overall running time.

CMPSC 465 Fall 2023 | Data Structures & Algorithms | Antonio Blanca and Yana Safonova

Quiz 7 solution

Recitation Section: **Date: Monday, Oct 16, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) For sparse graphs (graphs with few edges), Bellman-Ford can run asymptotically faster than Dijkstra's Algorithm.

- ◯ True
- ◯ False

Answer True
For a sparse graph with $|E| = O(1)$, Dijkstra is $O(|V|\log|V|)$ and Bellman-Ford is $O(|V|)$.

2. (2 pts.) While running Bellman-Ford on a graph with no negative cycles, updating every edge $|V|^2$ times instead of $|V|-1$ times has no effect on the output.

- ◯ True
- ◯ False

Answer True
The operation of updating the shortest path ending with a given edge is safe. Once the shortest paths are found, the output will not change.

3. (2 pts.) Which of the following recursions is used in Floyd-Warshall? Here, $P(i,j,k)$ is the shortest path from vertex i to vertex j using only the first k vertices and $w(i,j)$ is the weight of the edge from i to j.

- ◯ $P(i,j,k) = P(i,j,k-1) + w(k-1,j)$
- ◯ $P(i,j,k) = P(i,k,k) + P(k,j,k)$
- ◯ $P(i,j,k) = max\{P(i,j,k-1), P(i,k,k-1) + P(k,j,k-1)\}$
- ◯ $P(i,j,k) = min\{P(i,j,k-1), P(i,k,k-1) + P(k,j,k-1)\}$

Answer (d): $P(i,j,k) = min\{P(i,j,k-1),\ldots\}$
The new shortest path is the minimum between the known shortest path and some possible shorter path that goes through node k.

4. (2 pts.) Select **all** of the constraints a valid flow function $f(e)$ must follow (where e is an arbitrary edge).

- ☐ The sum of every $f(e)$ must equal the total flow leaving the source vertex.
- ☐ $f(e)$ must not exceed the capacity of e.
- ☐ The sum of the flow into a vertex must equal the sum of the flow out of that vertex.
- ☐ If e is part of a cycle, $f(e)$ must be zero.

Answer (b) and (c)
(b) is the capacity constraint and (c) is the flow conservation constraint. (a) is false because the sum of the flow on every edge includes the flow leaving the source vertex. (d) is false, here's a counterexample: $S \to T \to A \to S$. $S \to T$ is technically in the cycle, but flow can be routed through it.

5. (2 pts.) Every edge (u,v) in a flow network is converted into either a forward edge (u,v) or a backward edge (v,u) in Ford-Fulkerson's residual graph. (u,v) and (v,u) never exist simultaneously in the residual graph.

- ◯ True
- ◯ False

Answer False
If an edge in the flow network has some flow routed through it that is greater than zero but less than the capacity of the edge, it will have both a forward and a backward edge in the residual graph.

CMPSC 465 Fall 2023 | Data Structures & Algorithms, Antonio Blanca and Yana Safonova

Quiz 8 solution

Recitation Section: **Date: Monday, Oct 23, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) Which of the following problems has an optimal greedy solution?

- ◯ 0-1 Knapsack
- ◯ Minimum Spanning Tree
- ◯ Maximum sum path from root to leaf in a binary tree.

Answer Minimum Spanning Tree
Both Kruskal's Algorithm and Prim's Algorithm are optimal and greedy.

2. (2 pts.) You have developed two algorithms for the same problem. Algorithm A always produces an optimal solution in $O(2^n)$. Algorithm B is a greedy algorithm that runs in $O(n)$ but may not produce an optimal solution. Which of the following is true?

- ◯ Algorithm A is too slow to ever use.
- ◯ Algorithm B is too imprecise to ever use.
- ◯ Both algorithms have a use case.

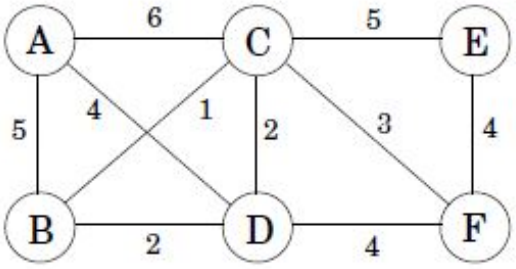
Answer Both algorithms have a use case.
A may be usable if n is small. B may provide a useful approximation of the optimal solution for applications where speed is more important than precision.

3. (2 pts.) Given some cycle C in an undirected graph G, the heaviest edge (u,v) in C may be included in the minimum spanning tree of G as long as there is a heavier edge on some other path between u and v in G.

- ◯ True
- ◯ False

Answer False
The heaviest edge on any cycle C can never be included in the minimum spanning tree. Alternate paths from u to v do not change this fact.

4. (2 pts.) What is the cost of the minimum spanning tree in the following graph?



- ◯ 12
- ◯ 14
- ◯ 15
- ◯ 16

Answer 14
See the solution to question 5 for a set of edges.

5. (2 pts.) Consider the graph in question 4. When running Kruskal's algorithm, which of the following is a valid order the edges can be added to the minimum spanning tree?

- ◯ (B-C), (B-D), (C-F), (A-D), (E-F)
- ◯ (C-D), (B-C), (C-F), (E-F), (A-D)
- ◯ (B-C), (B-D), (C-D), (C-F), (A-D)
- ◯ (B-C), (C-D), (C-F), (A-C), (C-E)

Answer (B-C), (B-D), (C-F), (A-D), (E-F)
These edges are discovered in increasing order of weight and have no cycles. Option (b) begins with an edge that is not the lightest. Option (c) forms a cycle between nodes B, C, and D. Option (d) ends with two sub-optimal edges.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 9

Recitation Section: **Date: Monday, Oct 30, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) The fully-optimized Disjoint Set implementation has a better asymptotic running time than the naive version for both the "union" and "find" operations.

- ◯ True
- ◯ False

Answer False
The naive implementation shown in class can perform the "find" operation in $O(1)$, while the optimal implementation's "find" is $O(\log n)$. "Union" does get faster, from $O(n)$ to $O(\log n)$.

2. (2 pts.) Huffman encoding does which of the following?

- ◯ Encodes the input data
- ◯ Compresses the input data
- ◯ All of the above

Answer All of the above
Huffman encoding both rewrites the input in a new alphabet (encoding) and reduces the number of bits the input string takes up (compression).

3. (2 pts.) In Huffman encoding, the symbol with the maximum frequency gets the fewest bits.

- ◯ True
- ◯ False

Answer True
To reduce the total space usage of the input string, the most common symbol should be encoded using the fewest bits.

4. (2 pts.) Which of the following is true about Huffman encoding?

- ◯ No code is the prefix of any other code
- ◯ The encoding is lossy
- ◯ Each input's encoding is unique
- ◯ Options (a) and (c) are both correct.

Answer No code is the prefix of any other code
Codes that match other codes' prefixes would cause ambiguity. The encoding is lossless (can be restored to its original alphabet), and an input with equal frequencies can produce several encodings.

5. (2 pts.) Using Huffman encoding, how many bits (0 or 1) are needed to store the 10-symbol string represented by the following table of symbols and frequencies?

Symbol	Frequency
C	5
M	2
P	1
S	2

- ◯ 20
- ◯ 22
- ◯ 18
- ◯ 16

Answer 18
An example encoding is:
C: 0, M: 10, P: 110, S: 111
Which gives bit count $5 \cdot 1 + 2 \cdot 2 + 1 \cdot 3 + 2 \cdot 3 = 18$

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 10 solution

Recitation Section: **Date: Monday, Nov 6, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) Given the Set Cover instance with items $B = \{1,2,3,4,5\}$ and sets $S_1 = \{3,4,5\}$, $S_2 = \{2,4,5\}$, $S_3 = \{1,5\}$, $S_4 = \{1,3\}$, what is the optimal (minimum) number of sets needed to cover all the items?

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4

Answer 2
We can choose S_2 and S_4 to cover all the items.

2. (2 pts.) Consider the same Set Cover instance from question 1. Assuming ties are broken by choosing the smaller subscript (e.g. S_1 wins a tie with S_2) the greedy algorithm finds the optimal number of subsets.

- ◯ True
- ◯ False

Answer False
Because of the tie-breaker, the greedy algorithm starts with S_1, then S_2, then S_3, for a total of 3 sets.

3. (2 pts.) Given an instance of the Set Cover problem with n items that must be covered, assume the optimal solution can cover all the items using k of the given sets. Which is the tightest valid upper bound on the number of sets the greedy algorithm uses?

- ◯ $\frac{n}{k}$
- ◯ k
- ◯ $k\ln(n)$
- ◯ kn

Answer $k\ln(n)$
The proof is too long to repeat here, see the lecture notes from Monday, Oct 30.

4. (2 pts.) Assume n and k are defined as in question 3. Pick the tightest valid lower bound for the size of the first set the greedy algorithm chooses to use.

- ◯ $\frac{n}{k}$
- ◯ k
- ◯ $k\ln(n)$
- ◯ kn

Answer $\frac{n}{k}$
The optimal solution uses sets k sets to cover n items, so at least one set must cover $\geq \frac{n}{k}$ items. The greedy algorithm can pick that set.

5. (2 pts.) The Longest Increasing Subsequence problem has an $O(n)$ optimal greedy algorithm.

- ◯ True
- ◯ False

Answer False
The $O(n)$ greedy algorithm from class is not optimal, which motivates the real optimal solution using Dynamic Programming.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 11 solution

Recitation Section: **Date: Monday, Nov 13, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) Let G be the underlying DAG of overlapping subproblems in a problem. We may use any topological order of G as an ordering for solving the subproblems to arrive at the global solution.

- ◯ True
- ◯ False

Answer True
Edges represent dependencies. Every topological ordering goes in an order where all prerequisites are resolved before the subproblems that depend on them.

2. (2 pts.) How long is the longest increasing subsequence in the following list of numbers?

7, 8, 1, 3, 9, 5, 6

- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

Answer 4
The longest increasing subsequence is 1, 3, 5, 6. Remember that it need not be contiguous.

3. (2 pts.) What is the edit distance between the strings "AGGT" and "GCT"?

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4

Answer 2
For example, insert "A" at the start of the second string, then change the second string's "C" to "G".

4. (2 pts.) Which of the following methods can be used to solve the Edit Distance problem? Select all that apply.

- ☐ Dynamic programming
- ☐ Recursion
- ☐ Greedy algorithm

Answer Dynamic programming and Recursion
The dynamic programming approach shown in class relies on a recurrence relation (as most DP solutions do), so it could be coded as a recursive function.

5. (2 pts.) Using the dynamic programming algorithm seen in class, which is the tightest upper bound on the running time to calculate the edit distance between two strings? Assume the longer string is length n.

- ◯ $O(\log n)$
- ◯ $O(n)$
- ◯ $O(n^2)$
- ◯ $O(n^3)$

Answer $O(n^2)$
It uses a table of size n^2, and each cell in the table takes $O(1)$ time to calculate.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 12 solution

Recitation Section: **Date: Monday, Nov 27, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) Because the chain matrix multiplication dynamic programming algorithm fills a table of size $O(n^2)$, its total running time is $O(n^2)$.

- ◯ True
- ◯ False

Answer False
The table size is correct, but each cell takes $O(n)$ to compute, so the total running time is $O(n^3)$.

2. (2 pts.) Which of the following is a valid objective function for a linear program?

- ◯ $5x^2$
- ◯ $x - xy$
- ◯ $5x + 2y - 3z$
- ◯ $\log(x+y)$

Answer $5x + 2y - 3z$
Only this function is linear (a sum of variables with constant coefficients). The incorrect options include exponents, variables multiplied by other variables, and logarithms, which are not linear.

3. (2 pts.) The point in the feasible region that optimizes the objective function of a linear program is always:

- ◯ On the boundary of the feasible region.
- ◯ Unique.
- ◯ Farthest from the origin.
- ◯ All of the above.

Answer On the boundary of the feasible region.
This is because the objective function is linear and the feasible region is convex (any line segment between two points is contained in the region).

4. (2 pts.) In linear program standard form #3, all of the decision variables must be non-negative, and all other constraints must be equalities.

- ◯ True
- ◯ False

Answer True
This is part of the definition of standard form.

5. (2 pts.) What is the maximum value of the objective function in the following linear program?

$$\begin{aligned} \text{maximize} \quad & 2x - y \\ \text{subject to} \quad & x + y \geq 1 \\ & x + y \leq 2 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

- ◯ 0
- ◯ 2
- ◯ 4
- ◯ Infinite (unbounded feasible region)

Answer 4
The objective function attains this value at $x = 2$, $y = 0$. To find this point, notice that we want to make x as large as possible while keeping y as small as possible. consider drawing the constraints as lines on the xy plane if you struggle to visualize them.

CMPSC 465 Fall 2023 — Data Structures & Algorithms — Antonio Blanca and Yana Safonova

Quiz 13 solution

Recitation Section: **Date: Monday, Dec 4, 2023**

Student Name: **PSU Email ID:**

1. (2 pts.) The initial values in the simplex method tableau always describe a feasible solution.

- ◯ True
- ◯ False

Answer False
With surplus variables (negative basic variables), the initial solution may not be feasible.

2. (2 pts.) The tableau form of the simplex method utilizes which of the following processes from linear algebra?

- ◯ LU factorization
- ◯ Gaussian elimination
- ◯ Matrix inversion
- ◯ Matrix multiplication

Answer Gaussian elimination
Gaussian elimination involves pivots, rescaling rows, and row elimination.

3. (2 pts.) The tableau form of the simplex method terminates when:

- ◯ All z row entries are non-negative.
- ◯ All basic variables have a non-negative coefficient.
- ◯ Either of the above becomes true.
- ◯ Both of the above are simultaneously true.

Answer Both of the above are simultaneously true.
If some basic variables are negative, the algorithm does not terminate at the normal z row non-negative condition.

4. (2 pts.) When some basic variable has a negative coefficient, we must choose the leftmost column with a positive entry in that row as the pivot column. Any other choice of column would give an incorrect result.

- ◯ True
- ◯ False

Answer False
Any column with a positive entry in that row will do. The leftmost one is chosen for consistency.

5. (2 pts.) Consider the following tableau as a final result of running the simplex method. Select the list with the optimal values of the objective function, x, and y.

	x	y	a_1	a_2	b
s_1	1	0	3	0	4
s_2	0	1	0	0	3
s_3	0	0	-2	1	7
z	0	2	5	3	14

- ◯ 24, 1, 1
- ◯ 2, 4, 3
- ◯ 14, 4, 3
- ◯ 7, 0, 2

Answer 14, 4, 3
The optimal objective function value is given by the bottom-right corner. The optimal values of x and y are the values in the b column corresponding to the row where that variable has a 1.